Willingness-to-Pay Utility Assessment: Feasibility of Use in Normative Patient Decision Support Systems

C. R. Flowers[1], A. M. Garber[1,2], M. R. Bergen[1], and L. A. Lenert[1]
[1]Stanford University School of Medicine, Stanford, CA, USA
[2]Palo Alto Veteran's Administration Medical Center, Palo Alto, CA, USA

The authors developed an automated patient interviewing tool to elicit individuals' willingness-to-pay (WTP) utilities under conditions of uncertainty and examined the reliability of this method and its potential usefulness in clinical decision support. We tested this method in 52 healthy volunteers using a computer-based interview that trained subjects in standard gamble (SG) and WTP methods, and elicited preferences for moderate Gaucher disease using WTP and SG. We assessed the validity of the WTP method by calculating the cost-effectiveness threshold implied by subjects' WTP and SG utilities; we also assessed subjects' understanding and comfort with using WTP for decision making by a questionnaire. The WTP method had good test-retest reliability (r = 0.796), and produced a cost-effectiveness ratio and ratings for understanding and clarity that support its validity. Moreover, many subjects felt that WTP was a reasonable (83%) method for therapeutic decision making and expressed comfort (62%) in using the method for their own health care decisions. These results suggest that a probabilistic method for WTP utility assessment is potentially useful for acquiring patient preferences for use in normative decision support systems.

INTRODUCTION

A long standing goal for researchers in Informatics has been to bring normative decision support to the bedside through access to computer-based decision models that provide recommendations that can be customized for individual patients[1]. When decision models are sensitive to patients' preferences regarding health outcomes, individual patient's utilities may need to be acquired in order to provide useful decision support[2]. One problem that has limited the usefulness of normative decision support systems has been the difficulty in obtaining utility values. Researchers inside and outside of the Medical Informatics community have investigated means to acquire utilities for normative decision making[3,4]. These techniques are broadly termed preference assessment methods, since typically an individual is asked to state preferences regarding one or more health outcomes. Although numerous methods exist for assessing utilities all methods have advantages and disadvantages.

The most well known methods for obtaining utility values are the Standard Reference Gamble (generally considered the index method) and the Time trade-off[4]. The Standard Gamble (SG) measures the significance of a health impairment to a patient using the risk of death that the patient is willing to accept to avoid life with the condition. The Time Trade-off measures the significance of an impairment in terms of the years of life expectancy that a person is willing to sacrifice to avoid life with the condition. Both of these methods for measuring utilities have been criticized for being too extreme to be realistic for use in clinical decisions[5]. When health impairments are minimal or last only a short period of time, accepting a risk of death or reducing life expectancy is not reasonable. For treatments that are low-risk and for diseases that do not reduce life expectancy the clinical relevance of the standard gamble and time trade-off may be limited.

An alternative method for obtaining patient utilities that are compatible with von Neuman-Morganstern utility theory is Willingness-to-Pay (WTP) assessment[6,7]. Considering the continued focus on the economics of health care, automated methods to assess the monetary values that patients attach to medical interventions may aid in therapeutic decision making. Using these data, normative medical expert systems that incorporate the principles of cost-benefit and cost-effectiveness analysis may provide means to help patients choose among health care alternatives.

Computer interviews have been used to acquired data regarding patients' health[8] and preferences for medical treatments[9-12] in a variety of settings. To generalize this approach, our research group developed the Interactive Multimedia Preference Assessment Construction Tool (IMPACT)[13], a tool for constructing preference elicitation interviews that uses multimedia methods to explain health states and measure preferences. This paper describes the development of software that extends IMPACT to assess preferences using WTP. The purpose of this study was to evaluate the feasibility of automated collection of WTP preferences for use in normative patient decision support systems and to compare the use of WTP to a previously evaluated SG method[9].

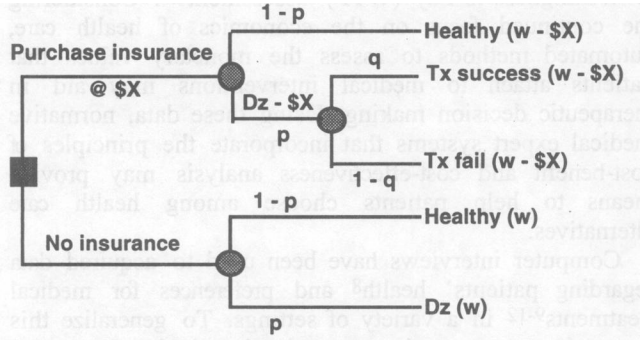
DESIGN CONSIDERATIONS

Willingness-to-pay methods measure the utility of an intervention in dollars rather than quality-adjusted life years (QALYs). Although WTP has been used to assess medical interventions[6,7], previous approaches have measured WTP under conditions of certainty. To be useful in normative decision support systems, WTP preferences should be obtained under conditions of uncertainty. The authors developed a probabilistic method that conditions WTP on two independent probabilities: the chance of contracting disease that requires treatment and the probability that the treatment is effective.

The first step in designing this instrument for assessing patients' WTP utilities was to devise a reasonable analogy for the probabilistic method. This WTP method was framed as a hypothetical insurance purchasing scenario. The decision tree that corresponds to this formulation of WTP is shown in Figure 1. This method of WTP assessment involves sequential decisions between purchasing insurance to cover the cost of treatment or not. The health state being assessed is associated with a given probability of requiring treatment (p in Figure 1). To simplify the decision presented to subjects, we described the treatment to be effective in all cases (i.e., $q = 1.0$).

The implementation of this analogy has three basic interface elements: a representation of the health state, a probability display, and a bidding game to find the subject's maximum WTP (Figure 2). Since the optimal method for performing WTP assessment is not known, we created a general tool that permits performing additional studies to explore how changing the utility elicitation interface influences patient preferences.

Figure 1. DECISION TREE FOR WTP UTILITY ASSESSMENT



X is the dollar amount being assessed in the WTP decision
p is the probability of developing symptomatic disease
q is the probability that the treatment eliminates the symptoms

IMPLEMENTATION

We implemented and tested a computer–based tool for the construction and delivery of WTP assessment interviews using Supercard 2.5 for the Macintosh OS. We created an editor environment for the WTP method to facilitate rapid construction and revisions of prototype interviews. To create a WTP assessment a researcher instantiates IMPACT objects to customize the assessment interface. Pictures and QuickTime movies can be incorporated to serve as reminders of the health states and the expense incurred for the treatment. Using the WTP editor a researcher creates the text that presents the assessment questions. As the text is displayed on the screen, a synchronized audio track is played, or a computer-generated voice reads the text using PlainTalk software (Apple Computers, Cupertino, CA). The researcher also sets the probabilities associated with the health state and selects a search method for obtaining the subject's maximum WTP. The automated interview can employ one of two methods for determining an individual's maximum WTP. The Titration method increases the WTP incrementally until it reaches the maximum WTP. The Ping-Pong method uses a search grid input by the researcher to determine the maximum WTP.

The WTP editor is internalized into the IMPACT program and it can be downloaded over the World Wide Web at http://preferences.stanford.edu. Using the editor, we created one instantiation of the WTP interface that assesses preferences for the treatment of moderate Gaucher disease, to compare with our previously developed SG method[11]. The text from the WTP assessment is shown below.

> If you had this disease in the way it was just described, what do you think that your quality of life would be? Suppose that a genetic counselor tells you that you have the gene for Gaucher disease. When you visit your doctor you discover that a highly effective treatment is available for Gaucher disease. However, this treatment is not covered by your current insurance. If you begin to experience symptoms from the disease, this treatment will restore you to perfect health. However, the treatment is quite expensive. Most people can only afford it if they purchase an insurance policy that covers the cost of the treatment. If you do not purchase the insurance policy you will not be able to receive the treatment if you need it. Like most insurance policies, this one will not be available to you once you begin to experience symptoms from the disease. When you ask your doctor for more information about your future you discover that 5 percent of people who have the gene for Gaucher disease get a form severe enough to eventually require treatment. That is, of 100 people like you who have the gene for Gaucher disease, 5 of them become ill enough to require medication to improve their health. The remaining 95 of them feel healthy enough to require no treatment at all. Think about the value of this medication to you. Consider how much you could realistically afford to pay each month. Would you be willing to pay at least X dollars per month for a policy that would cover the cost of the treatment if you ever needed it? If you would pay about this much, click on the button marked Pay this amount. If you would pay more than this amount, click on the button marked Pay more. If you would not pay this much money for an insurance policy to cover the cost of the treatment, click on the button marked Pay less.

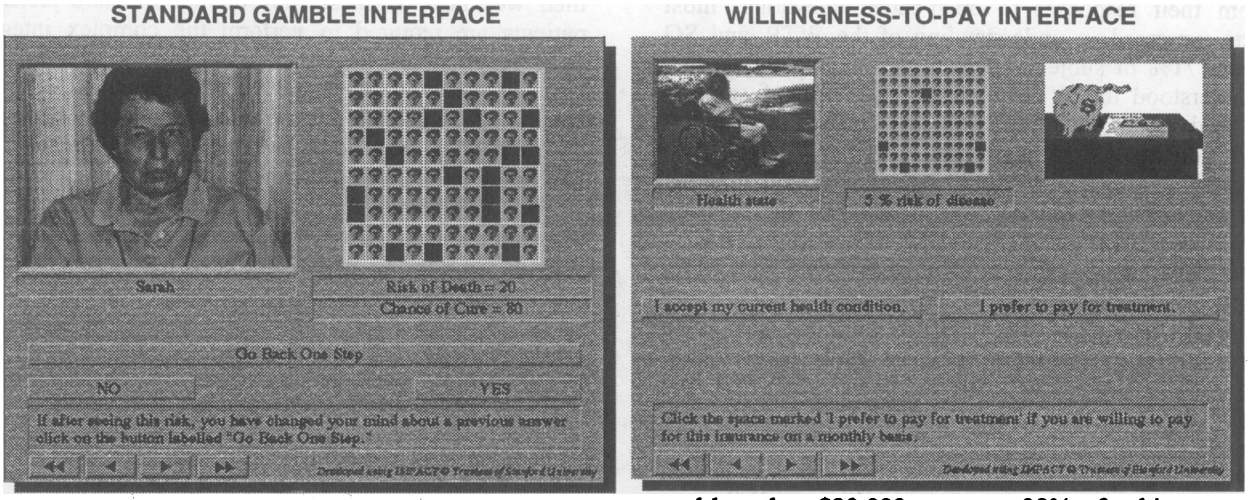
EVALUATION

Using the WTP and SG methods, the authors assessed preferences for a highly-effective treatment for Gaucher disease in a sample of 52 healthy volunteers who were recruited using advertisements and paid $10 for their participation in the study. This protocol trained subjects in SG and WTP methods, described a patient with moderate Gaucher disease, and then elicited utilities for this health state using the WTP and SG methods. QuickTime movies depicted a patient with Gaucher disease who experiences periodic, severe bone pain that is intermittently disabling. A full text of this health state description is in an article by Clarke and colleagues[11].

The interfaces for WTP and SG assessment are shown in Figure 2. The SG method interviewed subjects to determine the maximum risk of death that they would accept to avoid life with Gaucher disease. The SG used an animated panel of faces to display the risk of death (Figure 2). WTP was assessed as the greatest amount that a subject would pay per month to purchase an insurance policy to cover the cost of the treatment for Gaucher disease in the setting of a 5% probability of developing symptomatic disease.

The authors calculated mean SG and WTP preferences. We validated the reasonableness of WTP preference

Figure 2. SG AND WTP INTERVIEW INTERFACES



measurements with a questionnaire and by calculating subjects' implied cost-effectiveness threshold (in $ per QALY). This value is the coefficient, β_1, derived from regressing WTP values against the gain in QALYs from restoring subjects with Gaucher disease to normal health, $U_{WTP} = B_1 \cdot (1 - U_{SG})$. Subjects' implied cost-effectiveness threshold was examined to determine whether or not the WTP values assessed were consistent with meaningful valuation of one quality-adjusted year of life.

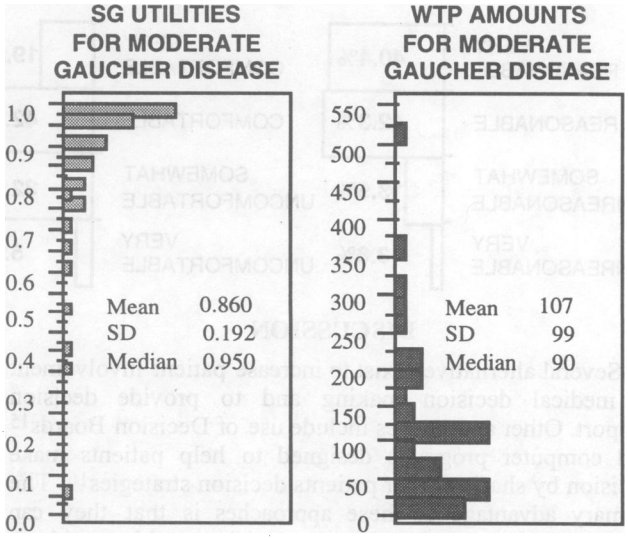
All subjects completed a questionnaire that examined their understanding of the WTP and SG methods using open-ended questions that told subjects to explain what they were asked to do during the SG and WTP assessments. A subject was judged to understand the SG method if she described the decision whether or not to accept a risk of death for access to the treatment. A subject was judged to understand the WTP method if he described the insurance purchasing scenario and the probability of becoming symptomatic. Answers were judged independently by two of the authors (CF and LL) and discussed to reach consensus. Subjects also rated the clarity, reasonableness, and difficulty of the SG and WTP methods using four-point discrete-valued scales. On a similar scale, subjects rated their comfort level using each to direct clinical decisions. Wilcoxon signed rank test was performed to detect differences in subjects' evaluations of each method. Subjects were retested two weeks after their initial visit, and Pearson's correlation was performed comparing test and retest SG and WTP values.

RESULTS

Subjects in the study were predominantly female (71%), single (60%), and without children (65%). 58% of subjects were between 18 and 30 years old, 17% were between 31 and 40 years, 12% were between 41 and 50 years, and 13% were older than 50. Four percent of subjects had earned a high school diploma only, 34% had attended 1 to 4 years of college, 33% had received a bachelor's degree, and 29% had attained a graduate or professional degree. 27% of subjects earned less than $20,000 per year, 38% of subjects earned between $20,000 and $40,000, and 35% of subjects earned more than $40,000.

Nearly all subjects, 98.1% (51 of 52), were willing to pay some amount of money for an insurance policy to cover the cost of the treatment for Gaucher disease, whereas 92.3% of the subjects (48 of 52) were willing to accept a risk of death in order to receive the same treatment. On average, subjects were willing to pay $107 ± $99 (S.D.; median $90) per month for insurance to cover the cost of treatment for Gaucher disease. SG and WTP utilities were not significantly different between test and retest. WTP and SG outlier values reflect the preferences of different individuals. The distribution of SG and WTP values is shown in Figure 3. Two-week test-retest correlation for WTP was comparable to other preference assessment methods ($r = 0.796$)[14]. Regression analysis of the retest WTP and SG values showed that subjects were willing to pay about $100,000 per QALY gained.

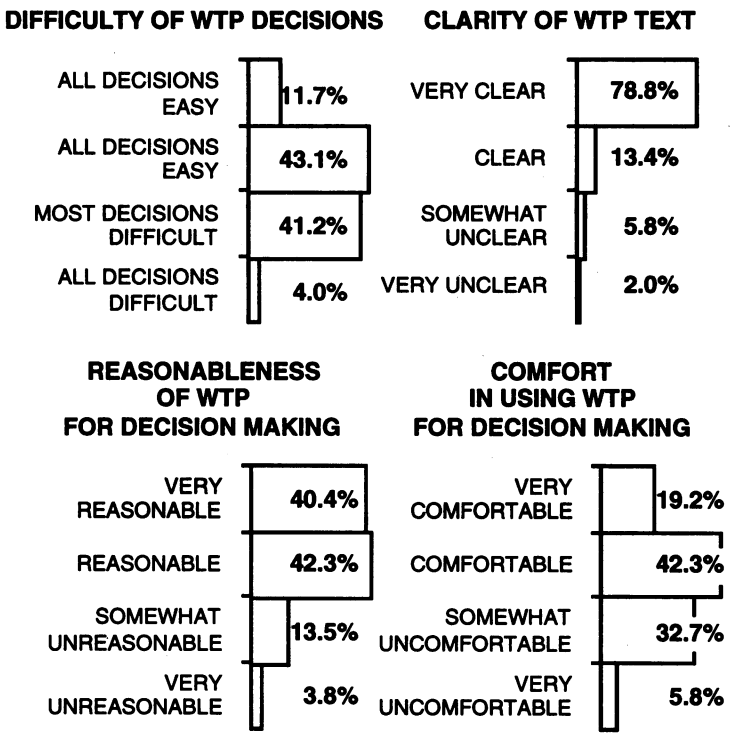
Figure 3. DISTRIBUTION OF SG AND WTP UTILITIES



From their responses to open-ended questions, most subjects expressed an understanding of the WTP and SG methods; 71% of subjects gave responses indicating that they understood the WTP method, and 79% of subjects understood the SG method. Subjects' ratings of the SG method indicated that they felt that it was very clear (85%) and moderately difficult (75%) to use in decision making. Eighty-one percent of subjects indicated that the decision that they were asked to make in the SG were reasonable. Fifty-seven percent of subjects were comfortable with using SG for decisions concerning their health care.

Subjects' ratings of the WTP method are shown in Figure 4. Subjects felt that the WTP method was very clear (79%), called for moderately difficult decisions (84%), and that the decision that they were asked to make were reasonable (83%). Furthermore, more than 61% of subjects felt comfortable or very comfortable with using this method for making decisions about their own health care. Wilcoxon signed rank tests showed no significant differences between the SG and WTP methods in terms of reasonableness, clarity, difficulty, comfort, and understanding ($p > 0.2$).

Figure 4. SUBJECTS' RATINGS OF WTP METHOD



DISCUSSION

Several alternatives exist to increase patient involvement in medical decision making and to provide decision support. Other approaches include use of Decision Boards[15] and computer programs designed to help patients make decision by sharing other patients decision strategies[16]. The primary advantage of these approaches is that they can increase patient involvement in decision making without the use of formal decision theory. However, herein also lies their weakness. Because they do not measure preferences, patients are required to perform the complex integrative tasks required to weigh risk and benefit. While patients typically reach an answer, there is no guarantee that their choice will be optimal from a decision theoretic perspective. The advantage of a normative decision support system is that potentially it can illustrate to patients the consequences of their preferences and by doing this help improve their decision making. However, use of computer models to provide feedback to patients from a normative perspective requires methods for measuring patient preferences.

In previous work we have shown that computer implementations of the Standard Gamble and the Time Trade-off can provide automated means to measure preferences[9,17]. In this paper, we demonstrate the extension of this work to Willingness-to-Pay utility assessment. On the whole, this implementation of the WTP method produced results on par with an existing computer-based SG method that has been shown in previous studies to be well-accepted in other clinical contexts[9]. The WTP method may be particularly well suited to measure preferences for conditions where it is unrealistic to trade-off life expectancy or to take a risk of immediate death (e.g., chronic conditions near normal health).

Previous work has focused on use of the WTP under conditions of certainty. However, if we are to use WTP preferences to represent a patient utilities for modeling decisions under conditions of uncertainty, we should assess those preferences under similar conditions. To provide a realistic scenario to assess WTP for health outcomes under conditions of uncertainty, we developed the insurance paradigm described in Implementation section. Results from testing this scenario in healthy volunteers suggest that it may be a useful method for measuring patients' preferences in WTP units. The method appeared to be reliable and produced cost-per-QALY estimates consistent with values described in recently published guidelines for cost-effectiveness analysis. Subjects responses to open-ended questions indicated that they understood the trade-off required by WTP assessment about as well as the SG. Furthermore, they felt that the tasks that they were asked to perform during the WTP method were at least, if not more reasonable than the SG. More than 61% of subjects felt comfortable using WTP to direct medical decisions--a rating that we feel strongly endorses the method.

The subjects in this study were young and well-educated and hence represent a near "ideal" population for use of computer-based decision support tools. Further studies of this interface need to be performed in older patients with lower education levels to evaluate the spectrum of patients in which WTP utility assessment would be useful. The text of the WTP assessment used in this study was not optimized for patients and reflects college-level reading skills. This level was appropriate for our study population, but before this method can be used in patients, the WTP decision task must be further simplified. While the population of this study was more educated than many

populations, its demographic features are similar to patients on the Internet[18]. This suggests that WTP methods may be useful in WWW-based normative decision support systems for patients such as the SecondOpinion program[19].

Conclusions and Future Work

Probabilistic WTP assessment appears to be practical to perform by automated computer interview, reliable, well understood, and potentially acceptable as a method to measure preferences for use in medical decision support systems. Currently, we are applying a similar methodology to develop of a World Wide Web interface for WTP assessment interviews. Future work will examine the acceptability of this method for measuring preferences for other conditions in patient populations. Ultimately, this method may facilitate acquiring preferences for health outcomes, and thus provide useful data for normative, patient decision support systems.

Acknowledgments

We thank our subjects and David Michelson, Roy Soetikno, and Manajeh Parineh. This work was supported by grants from the National Library of Medicine (LM-05626), and AHCPR (HS078180).

References

1. Pauker SG and Kassirer JP. Clinical decision analysis by personal computer. Arch Intern Med, 1981; 141(13):1831-7.
2. Heckerman DE, Horvitz EJ, and Nathwani BN. Toward normative expert systems: Part I. The Pathfinder project. Methods Inf Med, 1992; 31(2):90-105.
3. Farr BR and Schacter RD. Representation of preferences in decision support systems. Comp Biom Res, 1992; 25:324-335.
4. Torrance GW and Feeny D. Utilities and quality-adjusted life years. Int J Technol Assess Health Care, 1989; 5(4):559-75.
5. Mulley AG Jr. Assessing patients' utilities. Can the ends justify the means? Med Care, 1989; 27(3 Suppl):S269-81.
6. O'Brien B and Viramontes JL. Willingness to pay: a valid and reliable measure of health state preference? Med Decis Making, 1994; 14(3):289-97.
7. Gafni A. Willingness-to-pay as a measure of benefits. Relevant questions in the context of public decision making about health care programs. Med Care, 1991; 29(12):1246-52.
8. Slack WV, Safran C, Kowaloff HB, Pearce J, and Delbanco TL. A computer-administered health screening interview for hospital personnel. MD Comput, 1995; 12(1):25-30.
9. Lenert LA and Soetikno RM. Automated interviews to elicit utilities: potential applications in treatment of deep vein thrombosis. JAMIA, 1997; 4(1):49-56.
10. Sanders GD, et al. A computer-based interview to identify HIV risk behaviors and to assess patient preferences for HIV-related health states. Proc Annu Symp Comput Appl Med Care, 1994; :20-4.
11. Clarke AE, Goldstein MK, Michelson D, Garber AM, and Lenert LA. The effects of assessment method and respondent population on utilities elicited for Gaucher Disease. Journal of Quality of Life Research, in press.
12. Sumner W, Nease R, and Littenberg B. U-titer: a utility assessment tool. Proc Annu Symp Comput Appl Med Care, 1991; :701-5.
13. Lenert LA, Michelson D, Flowers C, and Bergen MR. IMPACT: an object-oriented graphical environment for construction of multimedia patient interviewing software. Proc Annu Symp Comput Appl Med Care, 1995; :319-23.
14. Froberg DG and Kane RL. Methodology for measuring health-state preferences--I: Measurement strategies. J Clin Epidemiol, 1989; 42(4):345-54.
15. Levine M, A G, Markham B, and MacFarlane D. A Bedside Decision Instrument To Elicit a Patient's Preference Concerning Adjuvant Chemotherapy for Breast Cancer. Annals of Internal Medicine, 1992; 117:53-58.
16. Barry MJ, Fowler FJ Jr., Mulley AG Jr., Henderson JV Jr., and Wennberg JE. Patient reactions to a program designed to facilitate patient participation in treatment decisions for benign prostatic hyperplasia. Med Care, 1995; 33(8):771-82.
17. Lenert LA, Morss S, Bergen MR, Goldstein MK, and Garber AM. Measurement of the validity of utility elicitation performed using a computerized interview. Med Care, in press.
18. Soetikno RM, Provenzale D, and Lenert LA. Studying ulcerative colitis over the World Wide Web. American Journal of Gastroenterology, in press.
19. Scott GC, Cher D, and Lenert LA. SecondOpinion: Interactive Web-based Access to a Decision Model. 1997 AMIA Fall Symposium, submitted.